

NBER WORKING PAPER SERIES

A MODEL OF THE SOCIALLY OPTIMAL
USE OF LIABILITY AND REGULATION

Steven Shavell

Working Paper No. 1220

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1983

The research reported here is part of the NBER's research program in Law and Economics. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

A Model of the Socially Optimal Use
of Liability and Regulation

ABSTRACT

Liability and safety regulation are examined as means of controlling risks in a theoretical model of the occurrence of accidents. According to the model, regulation does not result in appropriate reduction of risk -- due to the regulator's lack of knowledge about risk -- nor does liability result in that outcome -- because the incentives it creates are diluted by the chance that parties would not be sued for harm done or would not be able to pay fully for it. Thus, either liability could turn out to be superior to regulation or the reverse could be true. But as is stressed, joint use of the two means of controlling risk is generally socially advantageous, and the characteristics of their optimal joint use are determined.

Steven Shavell
Harvard Law School
Langdell 260
Cambridge, MA 02138
617-495-7920

A Model of the Socially Optimal Use of Liability and Regulation

S. Shavell*

I. Introduction and Summary

This paper will study a theoretical model of the occurrence of accidents in which the use of liability for harm done and of regulation of behavior will be examined as means of controlling accident risks. It will be assumed in the model that parties may, at a cost, take "care" to reduce accident risks, and that these risks may vary depending on their circumstances. It will therefore be socially desirable for parties to choose their levels of care in accordance with their particular circumstances; the greater the risks that they create, the higher should be their levels of care.

Neither regulation nor liability, however, will result in the model in parties taking the desirable, "first-best" levels of care. Specifically, regulation will not result in this outcome because the information about risk available to the regulatory authority will be imperfect. And liability will not lead to an ideal outcome because parties' incentives to take care will be inadequate (for reasons to be noted shortly), and possibly -- in an extension of the basic model to be studied -- because of incompleteness of their informa-

tion about risk. Depending on the importance of these factors, either liability or regulation could turn out to be preferred when they are considered as alternative methods for the control of risk, but (as will be stressed) it will often be socially advantageous for the two methods to be jointly employed.

basic model. The analysis of these points in the basic model will proceed as follows. It will first be asked how use of liability alone would function to reduce accident risks, supposing that parties have perfect knowledge of the risks they create but that their incentives to take care could be diluted for two reasons: their assets might not be sufficient to pay fully for harm done (meaning that they would not bear the full social cost of their actions); or they might not be sued for harm done. (These two causes of dilution of incentives are not unimportant:¹ it is frequently the case that a party's potential for doing harm is great in relation to its assets, even if the party is a large firm; and in some contexts, notably in regard to health-related and environmental risks, a party might not be sued for harm done because the harm might be difficult to trace to its source or because the harm might be highly dispersed.) Thus, it will indeed be the case that parties will choose to take less than the first-best levels of care; but because they recognize the magnitude of the risks they create, their levels of care will still generally bear a positive relation to the risks.

In the next section, it will be asked how use of regulation alone would work to control accident risks, presuming that parties must adhere to a regulatory standard of care as a precondition for engaging in their activities.² Further, this standard will be a uniform one, owing to an assumption that the regulator does not know parties' circumstances on an individual basis and therefore cannot modify the standard in view of them. Hence, the care parties take will not reflect the particular risks that they create; and under the (optimal) standard, parties presenting less than the average risk will have been forced to take more than their first-best levels of care, and those presenting more than the average risk will have been allowed to take less than their first-best levels of care.

The use of liability and regulation will then be considered as alternatives, and the comparison will depend on the factors just discussed: on the "informational" advantage of liability over regulation -- that decisions about care are made by those (the parties) possessing the better information about risk -- and on the relative disadvantage of liability -- that the factors which dilute incentives under liability are of no import where behavior is directly controlled.³

After making this comparison, it will be asked how joint use of liability and regulation would affect accident risks, where by joint use is meant that parties must satisfy the regulatory standard and will in addition be liable for harm done. Here it will be shown under a general condition

that many parties -- those creating other than relatively low risks -- will be led to do more than to satisfy the regulatory standard, for their potential liability will make that worth their while. And at the same time, just because such parties will take more care than is required, it will be socially desirable for the regulatory standard to be lower than if regulation were used alone; a reduction of the standard can in effect be afforded because liability is present to take up some of the "slack" associated with the lower standard. (Or at least liability can take up some of this slack if the incentive to take care is not too much diluted, which is the noted general condition.)

Last, joint use of liability and regulation will be compared to use alone of either of the two methods for the control of risk; and here under the general condition joint use will be best for the reasons just discussed; otherwise, however, joint use will offer no advantage.

extension of the model. An extension of the basic model will then be considered in which the regulator but not the parties know the value of a parameter that determines risk⁴ (along with parties' circumstances, which the regulator is still assumed to be unable to observe). This will raise the possibility that regulation can be used to correct for the systematically low level of care that would otherwise result if parties underestimated the parameter of risk.

numerical example; conclusion. Next, a numerical example that illustrates the results of the analysis will be presented, and following this, concluding comments will be made concerning the interpretation of the analysis and the importance of a factor that was omitted from it.

II. Basic Model

A. Description of model. Risk neutral parties may reduce the probability of causing harm by making expenditures on care, and the magnitude of the harm they might do will depend on their circumstances. Specifically, define

x = level of care; $x \geq 0$;

$p(x)$ = probability of causing an accident; $0 < p(x) < 1$;
 $p'(x) < 0$; $p''(x) > 0$; and

h = magnitude of harm if an accident occurs,

where h differs among parties, each of whom knows his own h . The regulator, however, is aware only of the distribution of h , and let

$f(h)$ = probability density of h ; $f(h) > 0$ on and only on $[a,b]$, $0 < a < b$.

Assume that the social welfare criterion is minimization of the expected sum of the costs of care and of harm done.

Thus, the first-best level of care as a function of a party's h is determined by minimizing over x

$$(1) x + p(x)h,$$

or by the first-order condition,⁵

$$(2) 1 = -p'(x)h,$$

which of course has the interpretation that the marginal cost of care, 1, equals the marginal benefits in terms of the expected reduction in harm, $-p'(x)h$. Denote by $x^*(h)$ the first-best level of care and observe that it is increasing in h .⁶

B. Liability as the sole means of controlling risk.

Assume that a party's potential for doing harm might exceed his assets; and assume also that the probability of being sued for harm done might be less than 1. Thus, define

y = party's assets, a non-negative amount; and

q = probability of suit given that harm has been done; $0 \leq q \leq 1$,

where y and q are the same for each party. Now if a party causes harm h and is sued, he will be liable for h ,⁷ but pay that amount only if $h \leq y$. Hence, his problem is to choose his level of care x to minimize

$$(3) x + p(x)q\min\{h, y\},$$

and let the solution to this be denoted

$$x_2(h) = \text{care selected by a party given } h.$$

Then we have

Proposition 1. Under liability, the care taken by parties as a function of the risk they present is

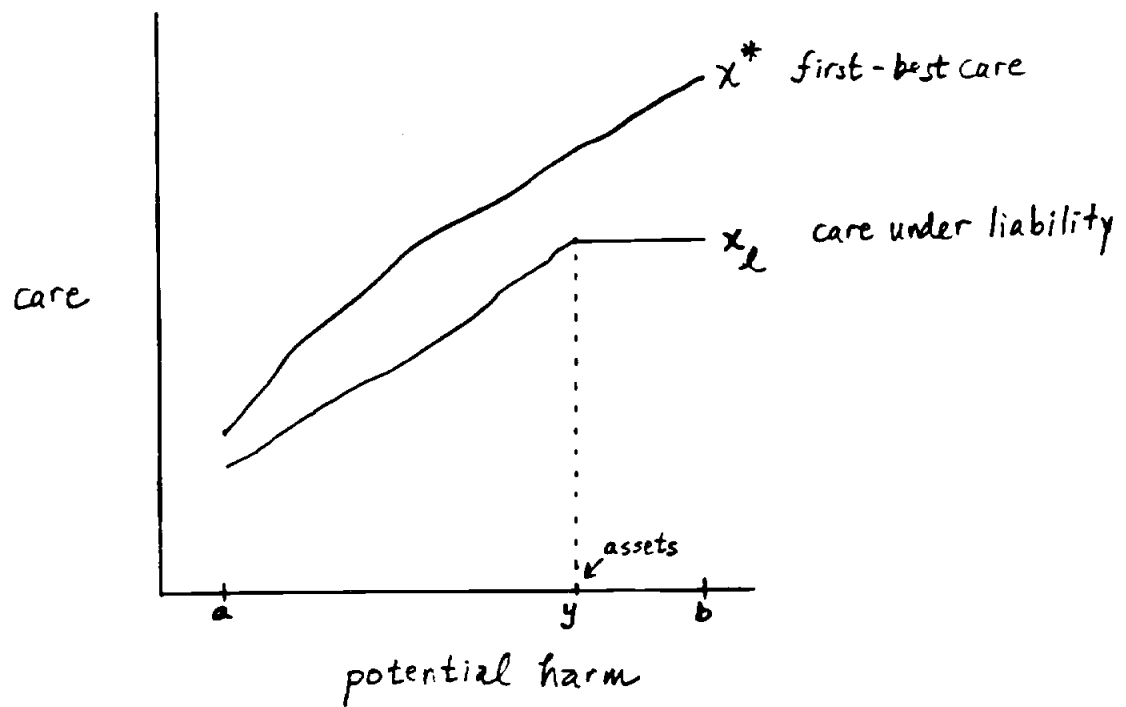
$$(4) \quad x_{\ell}(h) = \begin{cases} x^*(h) & \text{if } q = 1 \text{ and } h \leq y \\ q \min\{h, y\} & \text{otherwise;} \end{cases}$$

That is, unless the probability of suit is 1 and assets are sufficient to pay for harm, the incentive to take care will be diluted, and the level of care will be less than first-best, but it will increase with the magnitude of the risk so long as this is less than the level of assets.

Note. Graphs of x_{ℓ} and x^* (for a $q < 1$) are shown in Fig. 1.

Proof. Since the right-hand side of (3) is identical in form to (1), it is clear that $x_{\ell}(h)$ is determined by the first equality in (4). The inequality in (4) follows since (as observed above) x^* is increasing in its argument and since $q \min\{h, y\} < h$ if $q < 1$ or if $y < h$. Also, since $q \min\{h, y\} = qy$ for $h > y$, the shape of the graph of x_{ℓ} in Fig. 1 is explained. Q.E.D.

C. Regulation as the sole means of controlling risk. Because the regulator cannot observe h , the standard he sets must be uniform over the parties, and his problem of minimizing expected social costs is a simple one: If



First-best level of care vs. care taken under liability

Figure 1

s = regulatory standard,

and s must be satisfied in order for a party to engage in his activity, then the regulator's problem is to minimize over s

$$(5) \quad s + p(s) \int_a^b hf(h)dh = s + p(s)E(h),$$

where E denotes the expectation. Let

s^* = optimal regulatory standard,

the solution to (5). Then

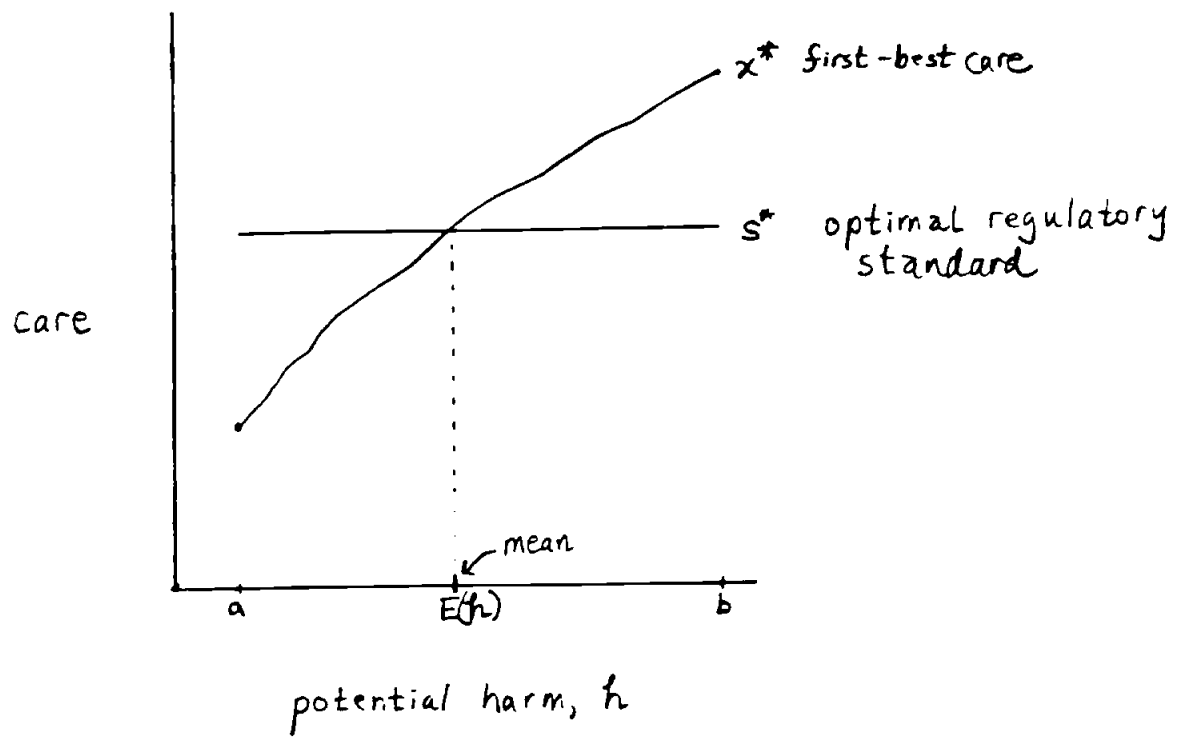
Proposition 2. Under regulation, the optimal standard equals the level of care that would be first-best for a party posing the average risk of harm, that is,

$$(6) \quad s^* = x^*(E(h));$$

in particular, parties presenting a risk h less than $E(h)$ take more care than is first-best, and those presenting a risk higher than $E(h)$ take less care than is first-best.

Note. Fig. 2 shows the situation under s^* .

Proof. Since the right-hand side of (5) is identical in form to (1), it is clear that s^* is determined by (6).
Q.E.D.



Optimal regulatory standard

Figure 2

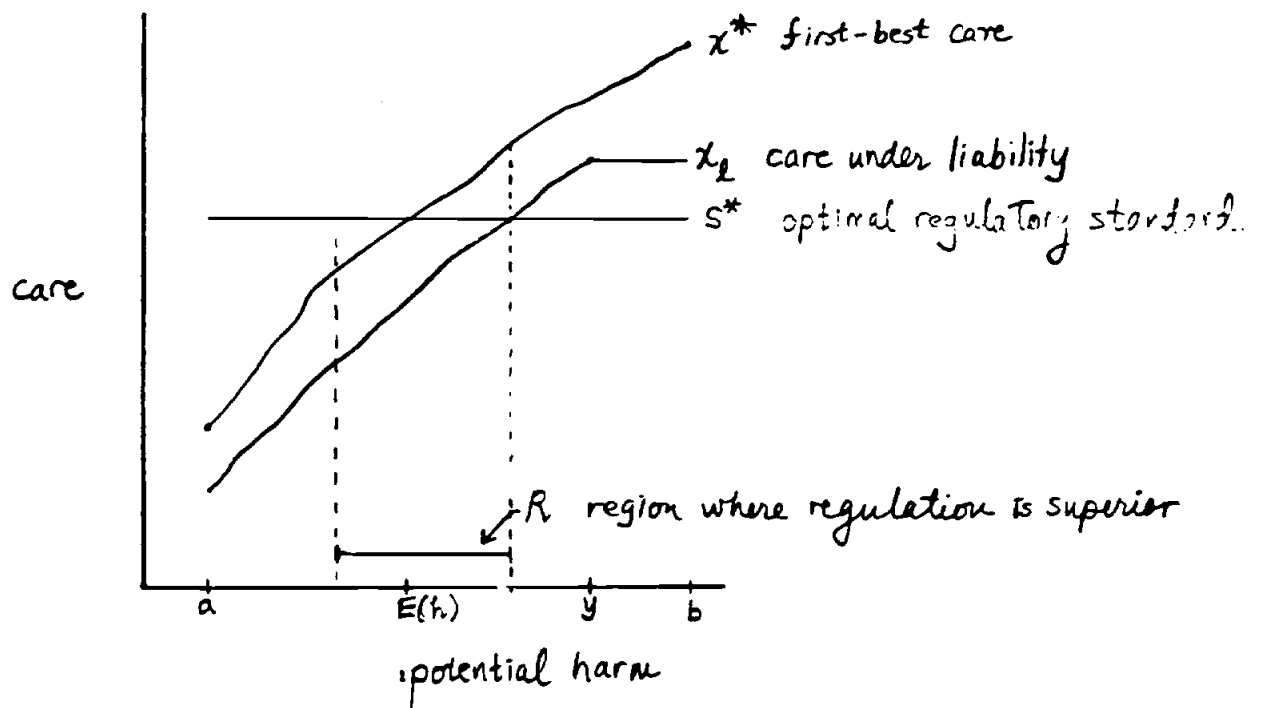
D. Regulation vs. liability. The difference in expected social costs between the situation where liability is employed and that where the optimal regulatory standard is used is

$$(7) \int_a^b \{ [x_\ell(h) + p(x_\ell(h)h)] - [s^* + p(s^*)h] \} f(h) dh.$$

From this, we obtain

Proposition 3. Use of regulation will be superior to use of liability if the factors that dilute the incentive to take care under liability are sufficiently important (q or y sufficiently low) or if the variability in different parties' circumstances is sufficiently small (h sufficiently concentrated about $E(h)$); otherwise liability will be superior to regulation.

Notes. More precisely, the claims are, first, that given y , there is a $q(y)$ where $0 < q(y) \leq 1$ such that regulation is superior to liability for $q \leq q(y)$ and liability is superior to regulation for $q > q(y)$. Second, that given q , there is a $y(q)$ where $0 < y(q) \leq b$ with properties analogous to those of $q(y)$. And third, that given $q < 1$ or $y < E(h)$, if the probability mass of h is sufficiently concentrated about $E(h)$, the regulation is superior to liability. To see why these claims should be true, consider Fig. 3, which shows that regulation will be superior to liability in a region R about $E(h)$ if x_ℓ lies below x^* . Now it is clear from the Figure that as q or y and therefore the graph of x_ℓ



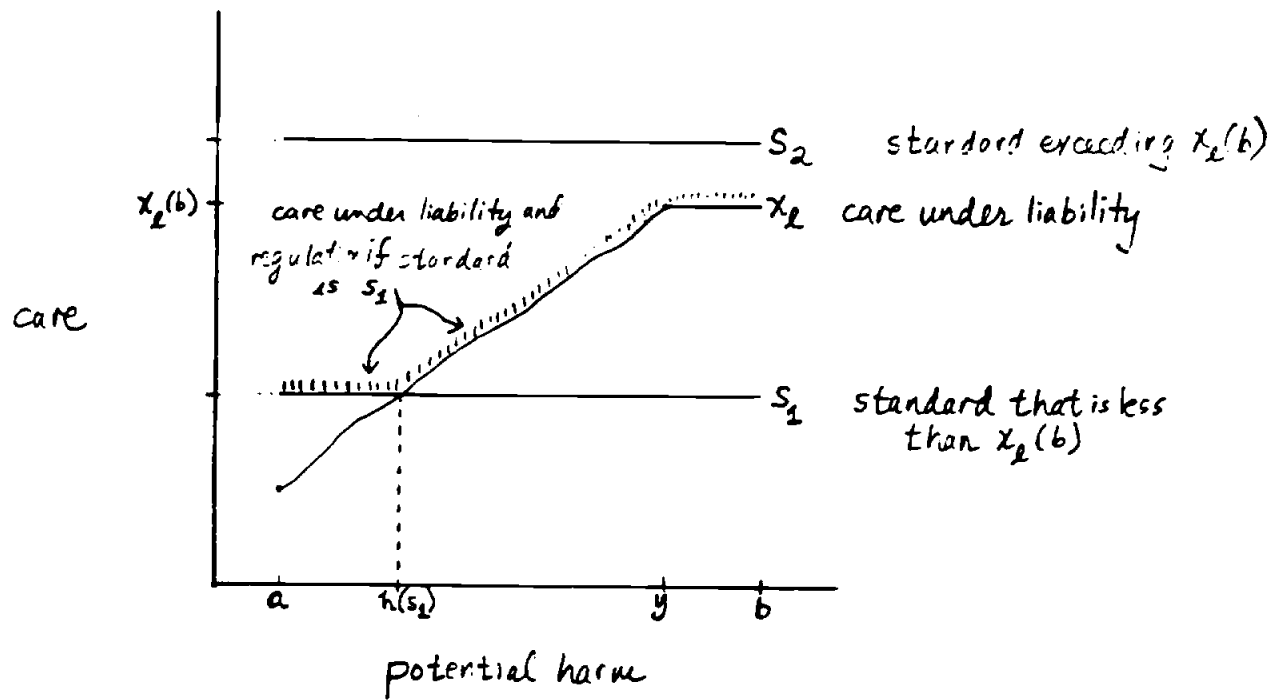
Comparison of regulation and liability

Figure 3

falls, the region R grows; and it is also clear that as the probability mass becomes more concentrated about $E(h)$, the likelihood of being in region R increases; thus the validity of the claims is evident.

Proof. If q or y equals 0, then (4) implies that $x_\ell(h)$ is identically equal to 0, and thus the situation is as if $s = 0$. But since s^* is the (unique) optimal s and is positive, social costs must be lower under regulation than when q or y equals 0. This fact and continuity of social costs in q implies that for any y , regulation is superior to liability for all q sufficiently small. Moreover, if liability is superior to regulation for some q_1 , then the same must be true for any $q_2 > q_1$; for social costs are easily shown to be decreasing in q under liability but are unaffected by q under regulation. The first claim therefore follows. The second claim is established similarly. Regarding the other claim, note from (4) that if $q < 1$ or $y < E(h)$, then $x_\ell(E(h)) < x^*(E(h))$, but the latter equals s^* . Hence, by continuity, there must exist a non-degenerate interval including $E(h)$ in which $s^* + p(s^*)h < x_\ell(h) + p(x_\ell)h$. Thus, if enough probability mass is within this interval, (7) will be positive, and regulation will be superior to liability. Q.E.D.

E. Joint use of regulation and liability. Assume now that parties must satisfy a regulatory standard and are also subject to liability. Then their levels of care will be given by $\max\{s, x_\ell(h)\}$, and the situation will be as shown in Fig. 4: For an s such as s_1 , all parties with $h \leq h(s_1)$



Joint use of regulation and liability

Figure 4

will take care of s_1 , and others will take care of $x_\ell(h)$.⁸ On the other hand, for an s such as s_2 that exceeds $x_\ell(b)$, all parties will take care of s_2 . With this in mind, it is evident that the problem of the regulatory authority is to minimize over s

$$(8) \int_b^a [\max\{s, x_\ell(h)\} + p(\max\{s, x_\ell(h)\})h]f(h)dh,$$

or, equivalently,

$$(8') \min_{0 \leq s < x_\ell(b)} \left\{ \min_{s \geq x_\ell(b)} \left\{ s + p(s) \int_a^b hf(h)dh \right\} + \int_{h(s)}^b [x_\ell(h) + p(x_\ell(h))h]f(h)dh \right\}.$$

Let s^{**} denote the solution to (8) and (8').⁹ Then we have

Proposition 4. Under optimal joint use of regulation and liability, there are two possible types of outcome in situations where the incentive to take care is diluted ($q < 1$, $y < b$):¹⁰

(a) The optimal regulatory standard is less than the optimal standard were regulation used alone, but it exceeds the first-best level of care for those parties posing the least risk; that is,

$$(9) x^*(a) < s^{**} < s^*;$$

furthermore, some parties are induced by liability to take more care than the required standard s^{**} . A sufficient condition for this case to hold is

$$(10) x_{\ell}(b) > s^*,$$

or, equivalently, that the incentive to take care is not too diluted (q sufficiently close to 1, y to b).

(b) Or, the optimal regulatory standard equals the optimal standard where regulation alone was employed; that is,

$$(11) s^{**} = s^*;$$

additionally, no party is induced by liability to take more care than s^{**} . This case will obtain if $x_{\ell}(b)$ is sufficiently low, or, equivalently, if the incentive to take care is sufficiently diluted (q or y sufficiently low).

However, in situations where the incentive to take care is complete ($q = 1$, $y \geq b$), then of course

(c) the optimal regulatory standard is 0.¹¹

Notes. (a) To explain this result and why s^{**} may be less than s^* (perhaps the typical case), consider Fig. 5a and the condition (10), which it is clear means that some parties are induced by liability to take more care than s^* . Now the reason that this condition implies that $s^{**} < s^*$ is essentially that when regulation alone was used, reducing the standard below s^* was not worthwhile because it resulted in all parties taking less care than s^* ; but here it results only in those parties with $h < h(s^*)$ taking less care -- for parties with higher h are induced by liability to take more care than s^* . Thus we said in the introduction that liability could take up some of the slack resulting from lowering the regulatory standard below s^* .

On the other hand, to understand why $s^{**} > x^*(a)$, observe that raising the standard from the level $x^*(a)$ would not lead to any first-order change in expected social costs in respect to parties with $h = a$ but would result in a first-order reduction in expected social costs in respect to all other parties with higher h who would not have been induced by liability to take as much care as $x^*(a)$.

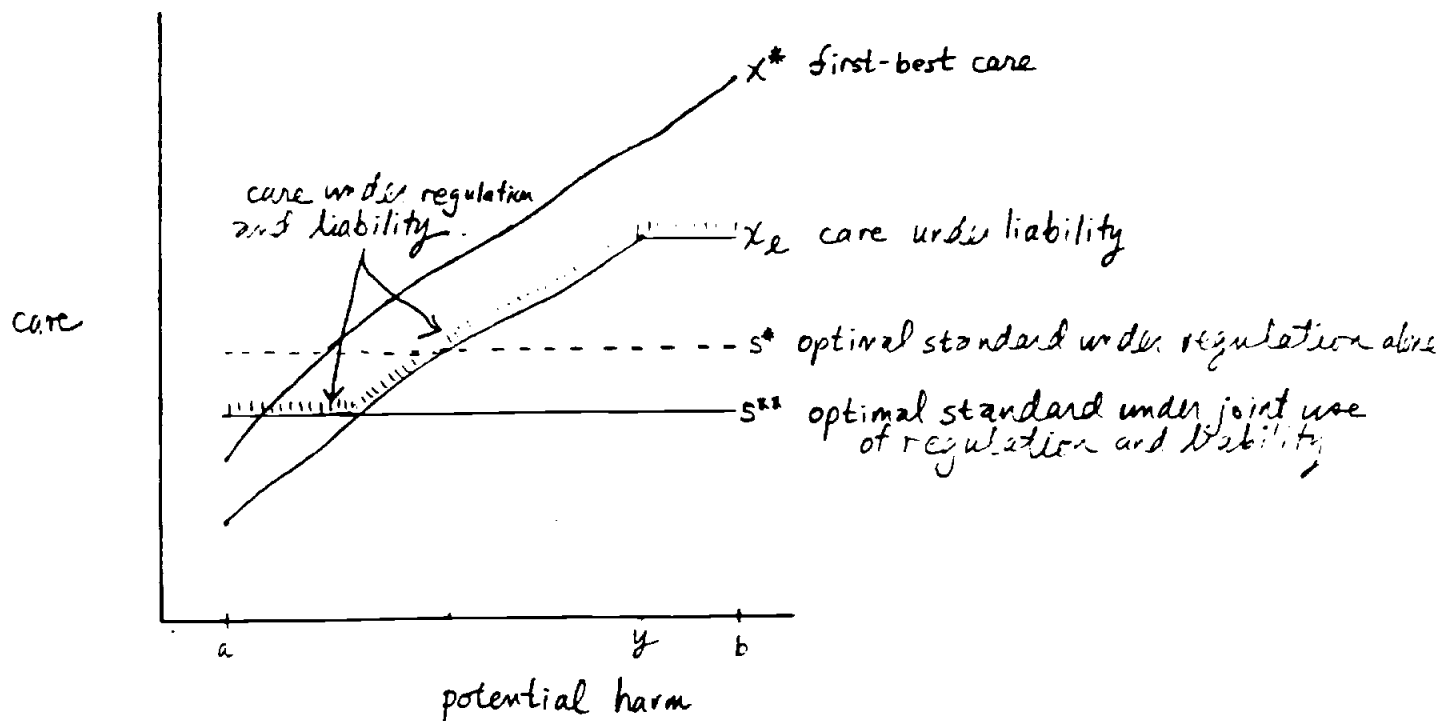
In addition, note from the proof that in the present case, s^{**} is determined by the condition

$$(12) \quad 1 = -p'(s) \left[\frac{\int_a^s h f(h) dh}{\int_a^s f(h) dh} \right],$$

the interpretation of which is that the marginal cost of care equals the expected reduction in harm, where the expectation is over only those parties who are not affected by liability and thus are affected by the regulatory standard.¹²

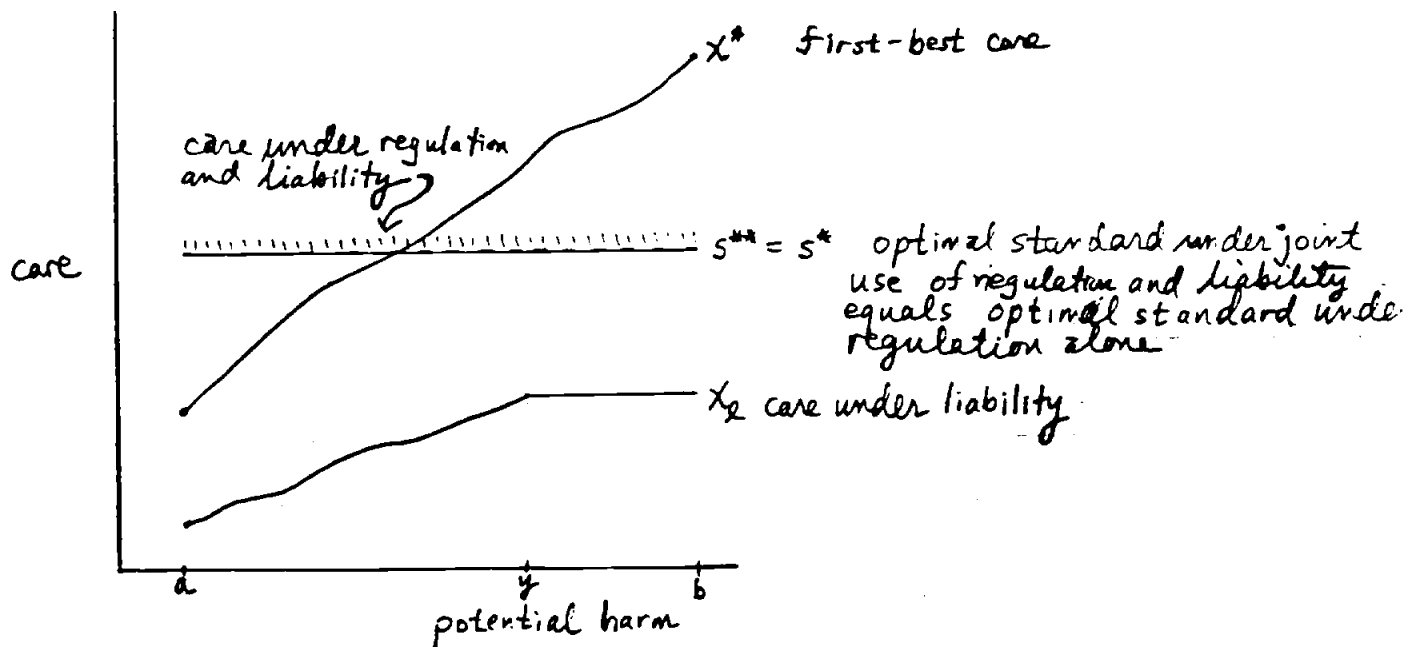
(b) It is evident from Fig. 5b why this case arises when $x_2(b)$ is sufficiently low: for then the incentive to take care created by liability is too weak to take up any of the slack due to lowering the standard below s^* ; it is therefore best to leave the standard at s^* .

(c) It is obvious that s^{**} should equal 0 in this case, since liability will induce each party to choose the first-best level of care.



Optimal joint use of regulation and liability: case where some parties are induced by liability to take more care than is required by regulation

Figure 5a



Optimal joint use of regulation and liability: case where all parties take only the required level of care

Figure 5b

Proof. See Appendix.

F. Joint use of regulation and liability vs. use of regulation alone or of liability alone. Finally, we may compare expected social costs resulting under joint use of regulation and liability (assuming $s = s^{**}$), under use of regulation alone (assuming $s = s^*$), and under use of liability alone; we then have

Proposition 5. Any of the three methods for the control of risk could turn out to be best:

(a) If the incentive to take care is sufficiently diluted (q or y sufficiently low), use of regulation alone is optimal; joint use of regulation and liability offers no advantage.

(b) Otherwise, if the incentive to take care is diluted ($q < 1$, $y < b$), joint use of regulation and liability is optimal.

(c) If the incentive to take care under liability is complete ($q = 1$, $y \geq b$), use of liability alone is optimal.

Notes.(a) More precisely, the assertion is that if q or y is low enough that under joint use, $s^{**} = s^*$, then use of regulation alone is equivalent to joint use of regulation and liability and is superior to use of liability alone. To explain this, recall from Proposition 4 that if $s^{**} = s^*$, then under joint use no party does more than take care of s^* . Thus the outcome is what it is under regulation alone.

And the outcome is different from, and therefore superior to, the outcome under liability alone (since $x_2(a) < s^{**}$).

(b) Here the assumption is that q and y are not so low that $s^{**} = s^*$. Thus we know from Proposition 4 that $x^*(a) < s^{**} < s^*$. This implies that the outcome under joint use of liability and regulation is different from that under liability alone (since $x_2(a) < s^{**}$) and from that under regulation alone (since $s^{**} < s^*$), and hence is superior to the outcomes under each.

(c) If $q = 1$ and $y \geq b$, then under liability alone a first-best outcome is achieved, so the result is obvious.

Proof. The Proposition is clear from the explanation just given in the Notes.

III. Extension of the model

Assume now that parties do not possess perfect information about risk; rather, they and the regulator each possess partial information of which the other is ignorant. Specifically, assume that the magnitude of harm should an accident occur is the sum of two elements: a random component depending on a party's particular circumstances and known by him but not by the regulator; and a parameter common to all parties but known only by the regulator. That is,

$$(13) \quad h = h_1 + h_2,$$

where

h_1 = component of harm depending on party's circumstances and known only by him,

h_2 = parameter determining harm known only by the regulator.

Also, let

$f(h_1)$ = probability density of h_1 ,

where as before f is positive on and only on $[a,b]$, and assume that the regulator knows the function f . Similarly, let

$g(h_2)$ = probability density of h_2 ; $g(h_2) > 0$ on and only on $[c,d]$, $0 < c < d$,

and assume that the parties know the function g . Note that given these definitions, the first-best level of care is still determined by minimizing (1), and thus by (2).

We will now reconsider what was done in the basic model. (In the interest of brevity, we will sometimes only sketch the results, as they (and their proofs) will be analogous to the previous ones.)

Under use of liability alone, the party's problem is to choose his level of care x to minimize

$$\begin{aligned} (14) \quad & x + p(x) \int_c^d q \min\{h_1 + h_2, y\} g(h_2) dh_2 \\ & = x + p(x) q E(\min\{h_1 + h_2, y\}), \end{aligned}$$

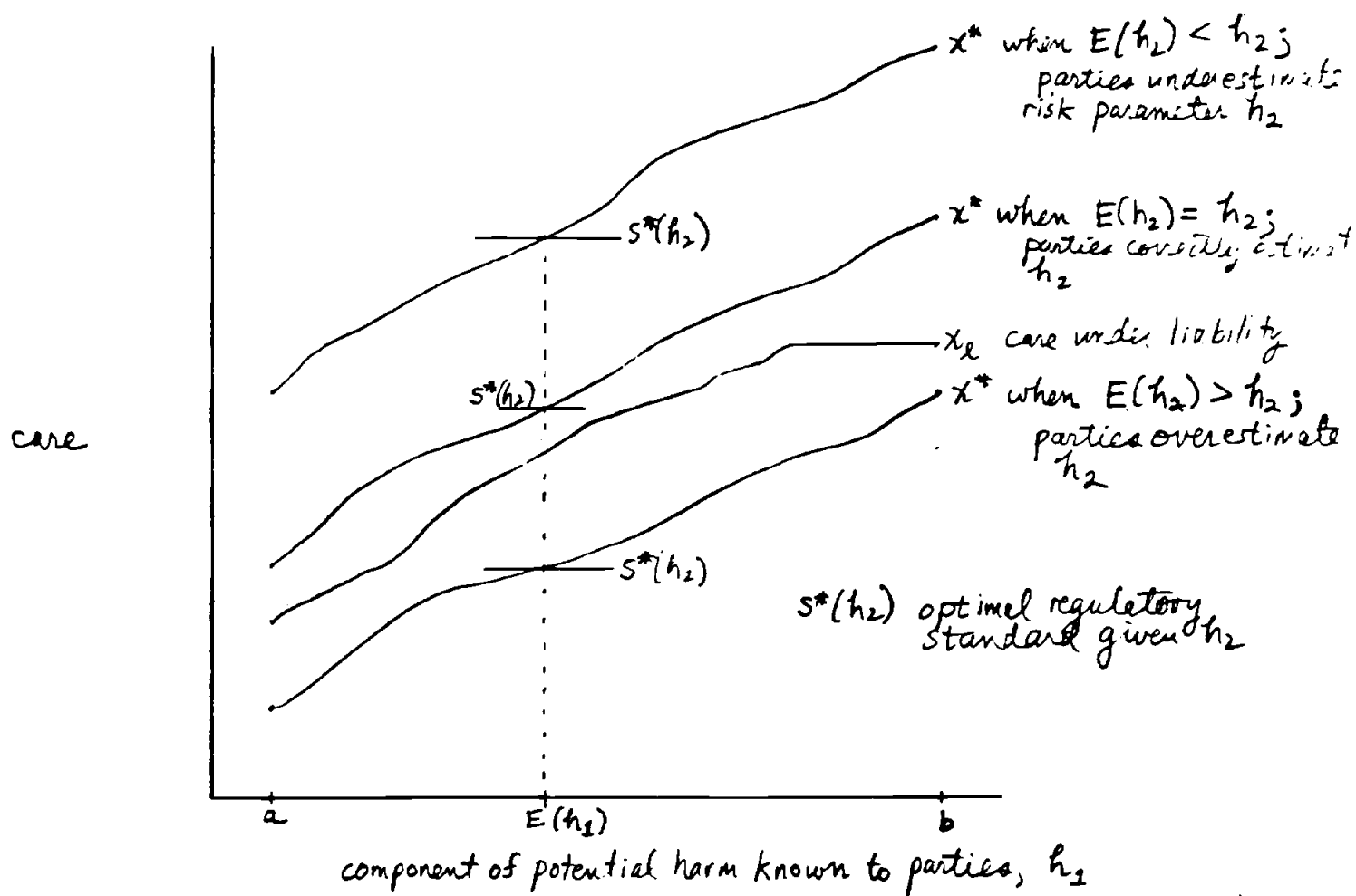
where it should be observed that the expectation here is over h_2 , since the party does not observe h_2 but does know h_1 . Let $x_\ell(h_1)$ denote the solution to (14). Thus the care taken by a party under liability is

$$(15) \quad x_\ell(h_1) = x^*(qE(\min\{h_1 + h_2, y\})).$$

It is clear from (15) that, similar to before, if $q < 1$ or if $y < h$ (with positive probability), then the incentive to take care will be diluted in the sense that $x_\ell(h_1) < x^*(h_1 + E(h_2))$. But unlike before, there is the additional element of parties' misperception of risk: If a party underestimates risk, that is, if $E(h_2) < h_2$, then the effect of dilution of incentives will be accentuated, and the gap between the party's care and first-best care will be made larger; whereas if the party overestimates risk, if $E(h_2) > h_2$, then the effect of dilution of incentives will be offset. This is illustrated in Fig. 6, where the graph of x_ℓ and the graphs of x^* are shown for three different values of h_2 .¹³

Under use of regulation alone, the problem of the regulator is to minimize over s

$$(16) \quad s + p(s) \int_a^b (h_1 + h_2) f(h_1) dh_1 = s + p(s) [E(h_1) + h_2],$$



Liability and regulation where parties' knowledge of risk is imperfect

Figure 6

where the expectation is over h_1 . If we let $s^*(h_2)$ denote the optimal regulatory standard given h_2 , then it is evident that

$$(17) \quad s^*(h_2) = x^*(E(h_1) + h_2).$$

Thus, as before, parties presenting less (more) risk than the average take more (less) care than is first-best; but now the standard depends on the information that the regulator alone possesses; it is higher the higher is h_2 , as is indicated in Fig. 6.

The difference in expected social costs between the situations where liability alone and where regulation alone is used to control risk is

$$(18) \quad \int_a^b [x_\ell(h_1+h_2) + p(x_\ell(h_1+h_2))(h_1+h_2)] - [s^*(h_2) + p(s^*(h_2))](h_1+h_2) f(h_1) dh_1,$$

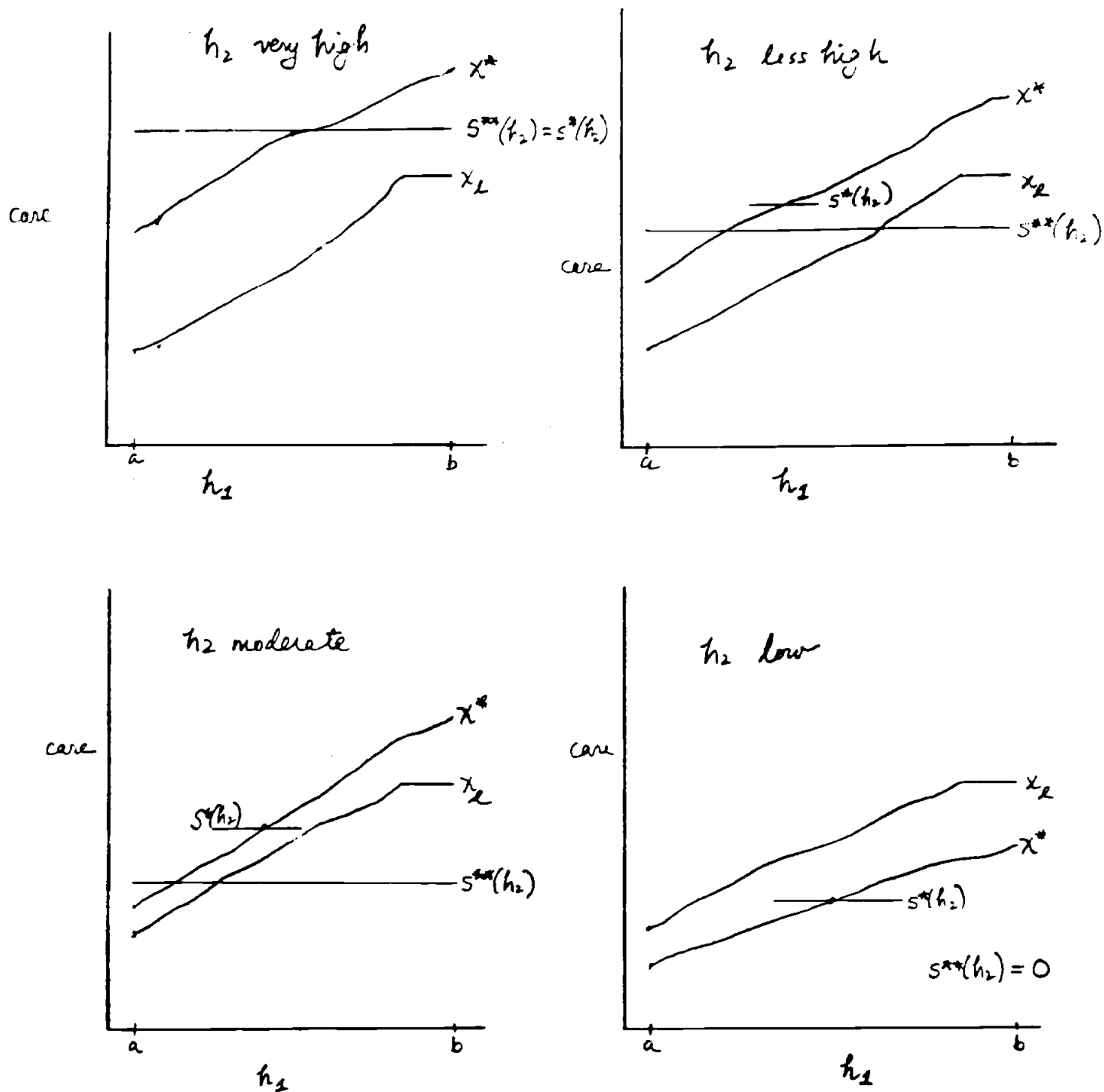
and it can be shown as before that for given h_2 , regulation will be superior to liability if q or y is sufficiently low or if h_1 is sufficiently concentrated about its mean; otherwise liability will be superior. Now, however, the comparison also depends on h_2 ; in particular, regulation will be superior to liability if h_2 is much larger than parties estimate, as is illustrated in the upper graph in Fig. 6.

Under joint use of regulation and liability, the problem of the regulator is to minimize the analog to (8'),¹⁴

$$(19) \quad \min \left\{ \min_{\substack{0 \leq s < x_\ell(b) \\ b}} \int_a^{h_1(s)} [s + p(s)(h_1 + h_2)] f(h_1) dh_1 \right. \\ \left. + \int_{h_1(s)}^{x_\ell(h_1)} [x_\ell(h_1) + p(x_\ell(h_1))(h_1 + h_2)] f(h_1) dh_1, \right. \\ \left. \min_{s \geq x_\ell(b)} s + p(s) \int_a^b (h_1 + h_2) f(h_1) dh_1 \right\},$$

and using this, the analog to Proposition 4 can be established: The optimal solution to (19), $s^{**}(h_2)$, may be less than $s^*(h_2)$ (the optimal solution under regulation alone), in which case some parties will take more care than $s^{**}(h_2)$; and a sufficient condition for this to hold is that $x_\ell(b) > s^*(h_2)$. Or, $s^{**}(h_2)$ may equal $s^*(h_2)$, in which case no party will take more care than $s^{**}(h_2)$. On the other hand, unlike in Proposition 4, h_2 affects s^{**} : For example, in Fig. 7a it is illustrated that if h_2 is sufficiently higher than parties' estimate of $E(h_2)$, then $s^{**}(h_2) = s^*(h_2)$; in Figs. 7b and 7c, it is shown that $s^{**}(h_2)$ declines as h_2 declines; and in Fig. 7d, it is shown that if h_2 is sufficiently lower than $E(h_2)$, then $s^{**}(h_2) = 0$.¹⁵

Last, comparing joint use of regulation and liability with use of either alone, then as before, any of the three alternatives could be best. Very briefly, and as can be explained along the lines of the Notes to Proposition 5, if



Optimal joint use of regulation and liability where parties' knowledge of risk is imperfect

Figure 7

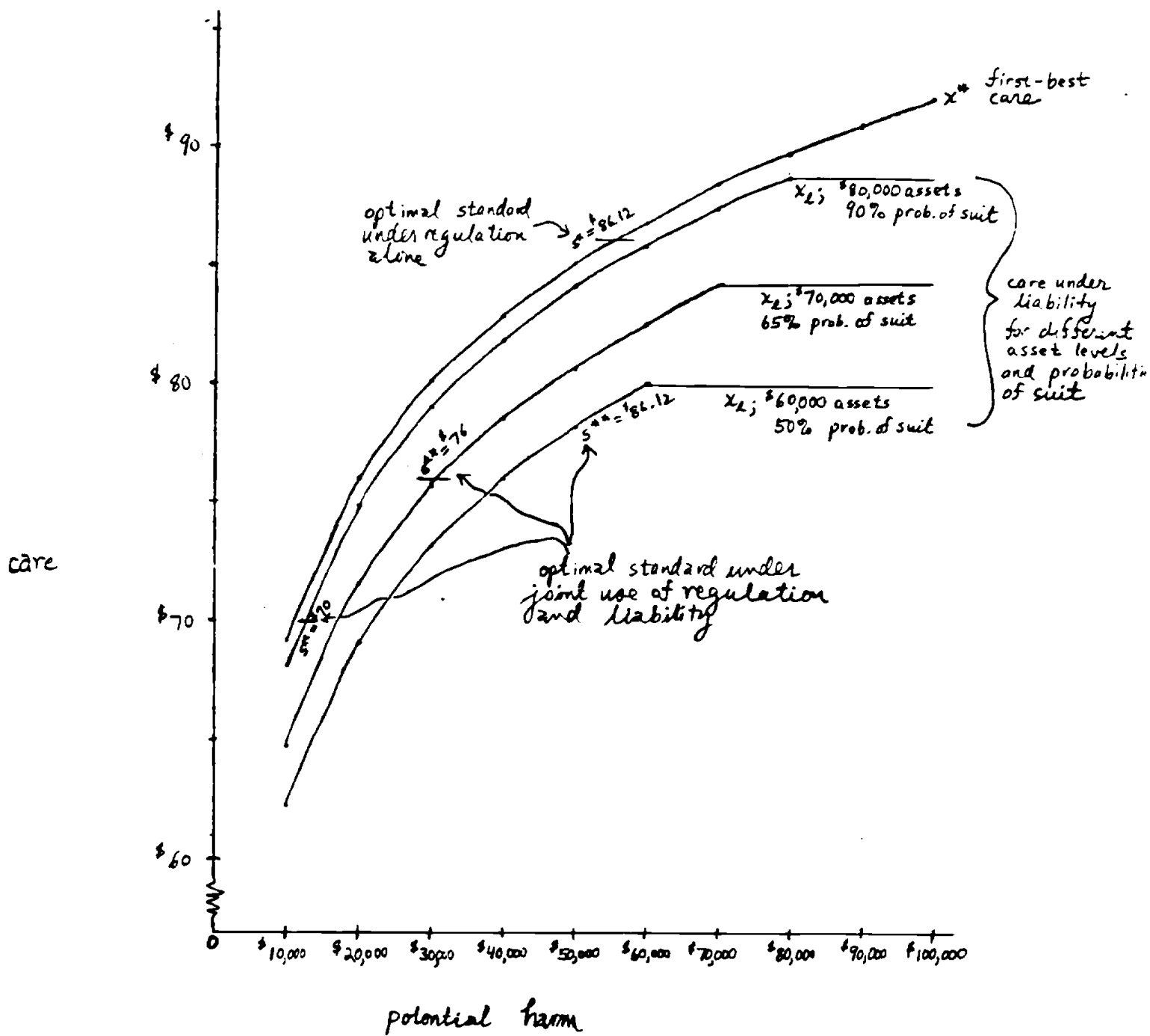
$s^{**}(h_2) = s^*(h_2)$, then use of regulation alone is optimal and joint use of it and liability offers no advantage. (This was the case in Fig. 7a, where h_2 was much greater than $E(h_2)$.) Second, if $x^*(a + h_2) < s^{**}(h_2) < s^*(h_2)$, then joint use of regulation and liability is optimal. (Figs. 7b, 7c.) Otherwise, use of liability alone is optimal and joint use of it and regulation offers no advantage. (This was the case in Fig. 7d, where h_2 was much less than $E(h_2)$.)

IV. Numerical example

Let us now present a numerical example illustrating the results of the basic model and then of its extension.

example of the basic model. According to the example, the probability of an accident as a function of expenditures on care is $p(x) = e^{-.1x}$; the harm that a party would cause if an accident occurred is known only by him; and this harm is equally likely to be any thousand dollar amount between \$10,000 and \$100,000 (that is, there are equal numbers of parties who would cause \$10,000 of harm, \$11,000 of harm, and so forth).

First-best expenditures on care¹⁶ for the example are shown in the upper graph of Fig. 8; and the expenditures that would be taken under liability alone¹⁷ are shown in the lower graphs for three different combinations of asset



Example of basic model

Figure 8

levels and probabilities of suit: \$80,000 and 90%; \$70,000 and 65%; and \$60,000 and 50%. Note that the first of these graphs becomes equal to a constant beyond \$80,000 of harm, for if a party's assets are \$80,000, the prospect of doing harm in excess of that amount will not cause him to take any more care than he would if the harm were limited to \$80,000; and similarly the other two graphs become constant beyond \$70,000 and \$60,000 respectively. Note in addition that these graphs are ordered in the expected way: as there is the least dilution of incentives when assets are \$80,000 and the probability of suit is 90%, the associated graph of care is the highest among the three, and as there is the most dilution when assets are \$60,000 and the probability is 50%, the associated graph of care is the lowest.

Also illustrated in Fig. 8 is the optimal standard of care (s^*) if regulation alone is employed. This standard equals \$86.12;¹⁸ if the standard is used, all parties who would cause less than \$55,000 of harm in an accident will have been required to spend more on care than is first-best, and all those who would cause more than \$55,000 of harm will have been allowed to spend less on care than is first-best.

Comparing the use of the \$86.12 regulatory standard to the use of liability alone, it turns out that use of liability is superior in the cases where assets and the probability of suit are \$80,000 and 90%, and \$70,000 and 65%; but use of the standard is superior in the case where assets

and the probability of suit are \$60,000 and 50%, owing to there being the greatest dilution of incentives in that case were liability employed.

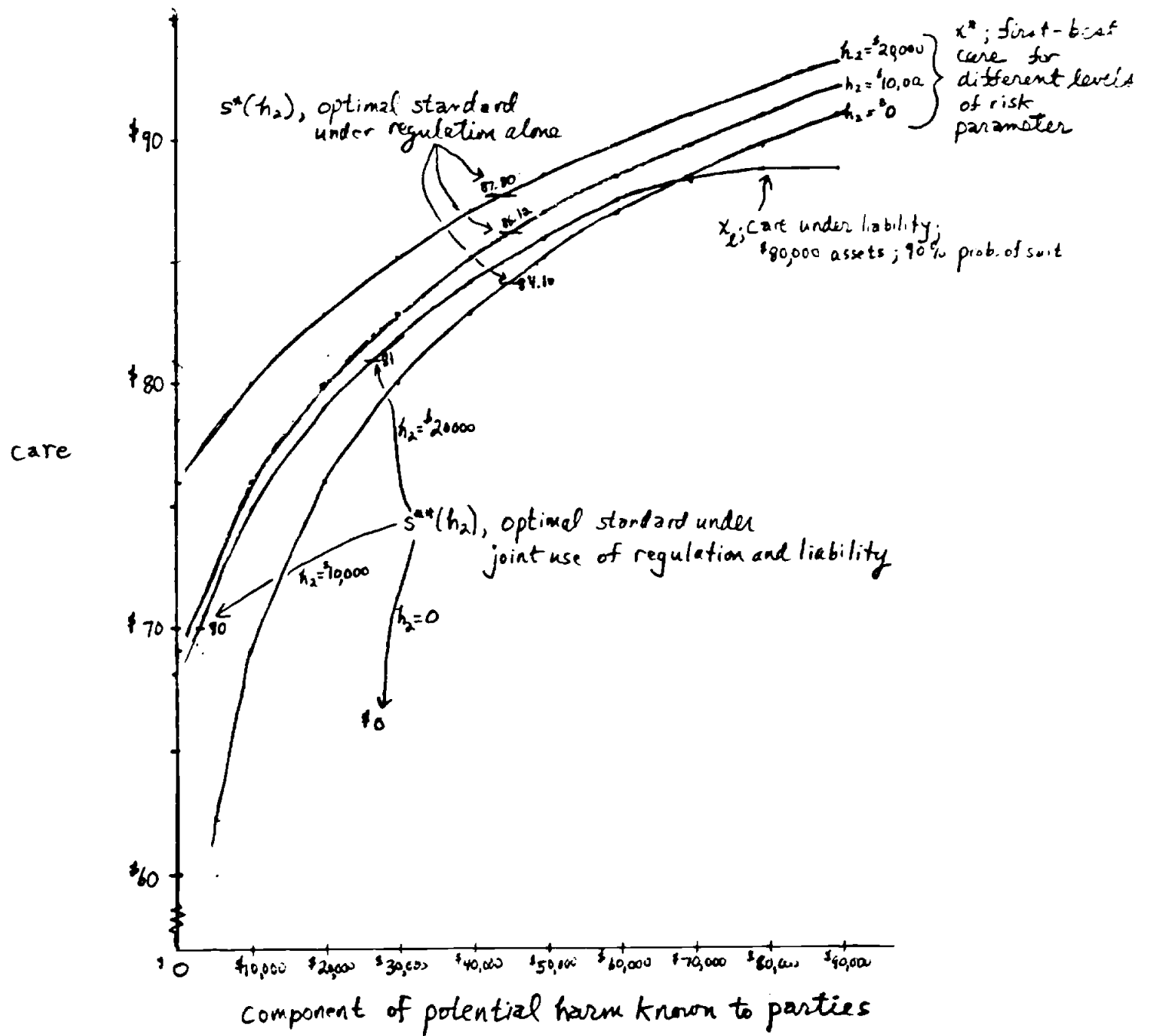
Finally, the optimal standards of care (s^{**}) assuming that regulation and liability are jointly employed are indicated in the Figure. In the case where parties' assets are \$80,000 and the probability of suit is 90%, the optimal standard is \$70; thus only a few parties (those who would cause less than about \$13,000 of harm) are actually affected by the standard; the rest spend more than \$70 on care due to their potential liability; the optimal standard of care is as low as it is for this very reason. Where parties' assets \$70,000 and the probability of suit is 65%, the optimal standard of care is \$76; now more parties are affected by the standard (all those who would cause less than about \$31,000 of harm); the standard is higher than in the last case because the dilution of incentives to take care due to liability was lesser in that case. And where parties' assets are only \$60,000 and the probability of suit is 50%, the optimal standard of care is \$86.12 -- the optimal standard if regulation alone is employed -- and all parties are affected by the standard. Therefore, and unlike in the other two cases, in this last case use of regulation alone is optimal; joint use of regulation and liability offers no advantage.

example of the extension of the model. Modify the example of the basic model by assuming that the harm that

parties would cause in the event of an accident is the sum of two elements: an amount known only by them and equally likely to be any thousand dollar amount between \$0 and \$90,000; and a parameter known only by the regulator and equally likely to be \$0, \$10,000 or \$20,000.

Graphs of the first-best levels of care as a function of the component of the potential harm observed by the parties are shown in Fig. 9 for each of the three different levels of the parameter known to the regulator;¹⁹ thus the graph where the parameter is \$20,000 is the highest and that where the parameter is \$0 is the lowest of the three.

Also shown in Fig. 9 is the graph of care that would be taken under liability alone if parties assets are \$80,000 and the probability of suit is 90%.²⁰ This graph is, of course, unaffected by the actual level of the parameter because it is known only to the regulator. The graph is somewhat below the first-best graph when the parameter is \$10,000, which is parties' estimate of the parameter, due to the dilution of incentives under liability; the graph is much below the first best graph when the parameter is \$20,000, for here the true risk is greater than parties estimate; and the graph is above the first-best graph (except over a small region) when the parameter is \$0, as in this case the true risk is less than parties estimate.²¹



Example of extension of the model

Figure 9

The optimal standard of care if regulation is used alone is shown in the Figure for each level of the parameter. The optimal standard rises as the parameter rises:²² it is \$84.12 when the parameter is \$0, \$86.12 when the parameter is \$10,000, and \$87.80 when the parameter is \$20,000. Use of these regulatory standards, however, turns out to be inferior to use of liability alone.

Last, the optimal standards of care if regulation and liability are used together are shown in the Figure. The optimal standard is \$70 if the parameter equals \$10,000 (note that \$70 is less than the \$86.12 standard were regulation alone employed); and the optimal standard is \$81 if the parameter is \$20,000 (and note that \$81 is less than the \$87.80 standard). In these two cases, joint use of regulation and liability is superior to use of regulation or of liability alone. But in the case where the parameter is \$0, the optimal standard is \$0, and it is not desirable to use regulation in addition to liability; the reason is that under liability alone, parties take more than enough care given the true, low risk.

V. Concluding Comments

Let us conclude with a brief consideration of the interpretation of the analysis and with a remark on a factor of importance that was omitted from it, namely, the magnitude of administrative costs.

(a) The choices that society evidently has made over use of liability and regulation may be identified, it is tentatively suggested, with the theoretically desirable choices of the model. That is, taking into account, on the one hand, the informational advantage (or lack thereof) of the private parties and, on the other hand, the possible sources of dilution of incentives to reduce risk, one can see a kind of rationality in the balance that has been struck between use of the two ways of controlling risk.²³

Consider first the area where society has placed chief reliance on liability as the legal means of controlling risk; consider, in other words, the familiar categories of accident coming under the rubric of tort (for example, I chop down a tree in my backyard and it strikes my neighbor's garage; I run for a bus and collide with someone). Here it does seem that as a general matter, the knowledge possessed by private parties about the particular dangers they create and the costs of meliorating them is significantly better than what would be enjoyed by a regulatory authority. (How much would such an authority know about the chance of this particular tree in my particular backyard damaging that particular neighbor's garage?) It also seems that the possibility that parties would not be able to pay fully for harm done does not represent a problem of great dimension (at least by comparison to the situation that often obtains in areas of regulation).²⁴ Similarly, it seems that a chance that parties will not face the threat of suit does not constitute

a substantial problem; for the usual tort, the harm done will not be difficult to attribute to the party who caused it (there will be no mystery about the source of the damage to my neighbor's garage or of the injury suffered by the person at the bus stop), nor will other reasons why suit might not be brought be of noticeable importance.²⁵

The situation appears different, however, in areas where use of regulation is significant. In such areas, the information that is necessary to justify regulation is sometimes virtually common knowledge (driving at high speed in a school zone or transporting explosives through tunnels obviously presents unacceptable risks) or at least it is fairly easy to obtain (appropriate procedures for pasteurization of milk); and where this is not true, the information is often of a special and technical nature that may be better appreciated by the regulatory authority. The latter seems at least plausible in regard to many of the environmental and health-related risks that are today increasingly subject to regulation, for expert medical, epidemiological, or ecological knowledge may be needed for their proper assessment. (Nevertheless, it is certainly not claimed that the informational advantage of the parties is uniformly small or nonexistent in regulated areas.)

Secondly, the risks in regulated areas can often readily be imagined to exceed the assets of responsible parties (hotel fires, the collapse of dams, and explosions, for example, can clearly result in losses surpassing the assets of the individuals or firms that cause them); thus the

incentives to reduce risk that are provided by liability alone may be seriously inadequate. This factor may be of particular importance again in respect to some of the health-related and environmental risks, as they may cause losses of extreme magnitude. (Consider the potential for harm associated with nuclear accidents, widespread exposure to carcinogens such as asbestos, mass use of drugs with harmful side-effects.) Furthermore, and still again in regard to many of the health-related and environmental risks, the incentives created by liability may be significantly diminished by the possibility that firms would not be sued for harm done; for such risks frequently result in harm that is difficult to trace to its source (was the individual's lung cancer due to conditions at his workplace, to medical x-radiation, or to "natural" causes?) or that occurs only after the passage of a long period of time; additionally, in some instances the harm is so dispersed that no single injured party would find legal action worthwhile.²⁶

The hypothesis that the actual use of regulation and liability is consistent with their desirable use in theory is given further support by the fact that many if not most risky activities are subject to both regulation and liability, and that according to general legal principle, satisfaction of regulatory requirements does not render parties immune from liability;²⁷ additionally, regulatory requirements seem to be regarded as minimal. These qualitative characteristics of the joint use of regulation and liability are, recall,

those derived in the analysis, where it was shown that it might well be advantageous to employ both means of controlling risk and to adopt a "low" regulatory standard (one less than the standard were regulation used alone) because the incentives created by liability would lead most parties to take a higher and tolerably good level of care given their circumstances.²⁸

The hypothesis of consistency of theory with fact should not of course be over interpreted. As stated at the outset, the consistency is thought to exist only in a very approximate sense, for the choices actually made concerning regulation and liability are undoubtedly influenced by a multitude of factors going beyond those of the analysis -- notably, the pressures of interest groups²⁹ -- and in any event often will not be the result of a conscious and explicit use of a social cost-benefit calculus. It is therefore hardly surprising sometimes to hear assertions that regulation is onerous or other times that it is lax.

(b) The principal normative implication of the analysis is, obviously, that in assessing how much to regulate and how much to rely on liability to control risk, careful attention should be paid to the locus of information about risk and to the importance of the factors that dilute the incentive to take care under liability.

(c) Regarding the omitted issue of administrative costs, the main point to be made is that there seems to be an underlying advantage in favor of liability due to its

being employed, by its nature, only if harm is done, whereas adherence to regulation is determined, and the associated administrative costs are borne, before, or at least independent of, the actual occurrence of harm. (Consider the extreme case where taking appropriate care completely eliminates the possibility of harm. Here there would be no administrative costs whatever under liability -- for if parties were induced to take appropriate care, there would never be an occasion for its use -- but the costs of regulation would be positive.³⁰) A fuller analysis of liability and regulation would recognize this apparent advantage of liability.³¹

Appendix

Proof of Proposition 4. The argument consists of a series of steps, the first four of which establish (a) and the last three of which show (b).

(i) s^{**} must lie in $[x^*(a), s^*]$: It is very easy to verify that for every h , expected social costs are lower at $s = x^*(a)$ than at smaller s , so that $s^{**} \geq x^*(a)$.

To demonstrate that $s^{**} \leq s^*$, let $C(s;r)$ be expected social costs given s when regulation alone is employed, and let $C(s;r_l)$ be expected social costs when regulation and liability are jointly employed. Then for any $s_1 < s_2$, we claim that

$$(A1) \ C(s_1;r) - C(s_2;r) \geq C(s_1;r_l) - C(s_2;r_l).$$

This can be demonstrated by showing that the corresponding weak inequality in social costs holds for every h ; that is,

$$(A2) \ [s_1 + p(s_1)h] - [s_2 + p(s_2)h] \geq \\ [\max\{s_1, x_2(h)\} + p(\max\{s_1, x_2(h)\})h] - \\ [\max\{s_2, x_2(h)\} + p(\max\{s_2, x_2(h)\})h].$$

To verify (A2), consider Fig. 10, which shows in the regions A, B, and C the different possible relations that may hold among s_1 , s_2 and $x_2(h)$. It is clear that for h in the region A, equality obtains, for the parties will act identically under regulation alone and under joint use of regulation and liability. For h in region B, (A2) will be strict: under regulation alone parties will increase care from s_1 to

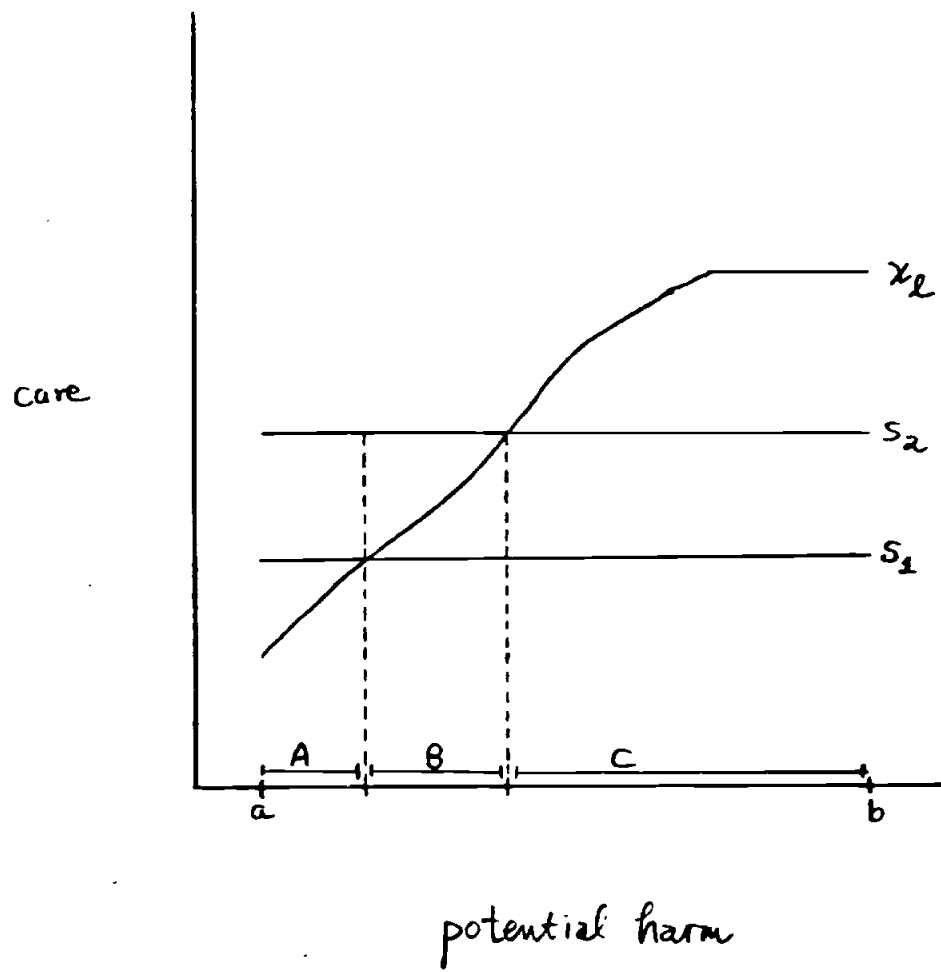


Figure 10

s_2 ; under joint use of regulation and liability, they will increase care only from $x_\ell(h)$ to s_2 ; (A2) thus becomes $s_1 + p(s_1)h \geq x_\ell(h) + p(x_\ell(h))h$; but the latter inequality holds strictly since expected social costs are convex in care and $s_1 < x_\ell(h) < x^*(h)$. For h in region C, (A2) will also be strict: under regulation alone parties will increase care from s_1 to s_2 ; since expected social costs are convex in care and $s_1 < s_2 < x^*(h)$, expected social costs are lower at s_2 than at s_1 ; but under joint use of regulation and liability, care will remain at the level $x_\ell(h)$, and costs will consequently be unchanged.

Having established (A2) and therefore (A1), suppose that $s^{**} > s^*$. Then since s^{**} minimizes $C(s; r_1)$ over s , we know in particular that $C(s^*; r_1) - C(s^{**}; r_1) \geq 0$. But (A1) then implies that $C(s^*; r) - C(s^{**}; r) \geq 0$, which contradicts the fact that s^* is the unique s that minimizes expected social costs under regulation alone. Hence it must be that $s^{**} \leq s^*$.

(ii) If $s^{**} < s^*$, then some parties will choose care exceeding s^{**} : Equivalently, we are to show that $x_\ell(b) > s^{**}$. Thus, suppose the contrary. Then for $s \geq s^{**}$, the second term in braces in (8') is relevant. Now since this term has a unique minimum over all s at s^* and $s^* > s^{**}$, the term must have a unique minimum over $s \geq s^{**}$ at s^* . But this means that the term cannot have had a minimum at s^{**} , which is a contradiction.

(iii) If $s^{**} < s^*$, then s^{**} is determined by the first-order condition (12) and $s^{**} > x^*(a)$: From (ii), we know that if $s^{**} < s^*$, then the first term in braces in (8') is relevant for all s in an interval properly including $x^*(a)$ and s^{**} . In particular, then, at s^{**} the derivative with respect to s of the term must be zero. But the derivative is

$$(A3) \quad \int_a^{h(s)} f(h)dh + p'(s) \int_a^{h(s)} hf(h)dh,$$

or, equivalently,

$$(A3') \quad \left\{ \int_a^{h(s)} f(h)dh \right\} \left\{ 1 + p'(s) \left[\frac{\int_a^{h(s)} hf(h)dh}{\int_a^{h(s)} f(h)dh} \right] \right\}.$$

Now the first term of (A3') is positive, since we know from (i) that $s^{**} \geq x^*(a)$, since $h(x^*(a)) > a$,³² and since $h(\cdot)$ is increasing in its argument. Thus, if (A3') equals zero, its second term must equal zero, which is (12).

To demonstrate that $s^{**} > x^*(a)$, we need only show that (A3') is unequal to zero, and in fact negative, when evaluated at $x^*(a)$. (This will prove the result since by the last paragraph it will mean that s^{**} is unequal to $x^*(a)$, and by (i), $s^{**} \geq x^*(a)$.) Observe first that since $h(x^*(a)) > a$, the first term in (A3') is positive. Notice second that

$$(A4) \quad \frac{\int_a^{h(s)} hf(h)dh}{\int_a^{h(s)} f(h)dh}$$

is the mean of h conditional on its being in the interval $[a, h(s)]$. The reader can easily verify that this tends to a as $h(s)$ tends to a and is strictly increasing in $h(s)$. Thus, since $h(x^*(a)) > a$, (A4) must exceed a at $s = x^*(a)$. And it follows from this and the fact that $1 + p'(x^*(a))a = 0$ that the second term in braces in (A3') is negative when evaluated at $x^*(a)$.

(iv) If (10) holds, then $s^{**} < s^*$: Suppose otherwise. Then by (i) $s^{**} = s^*$. But since (10) implies that the first expression in braces in (8') is relevant at s^* , we need only show that (A3') is positive at s^* to contradict the optimality of s^* . To do this, note from (6) and (2) that

$$(A5) \quad 1 + p'(s^*) \int_a^b h f(h) dh = 0.$$

However, since (A4) is strictly less than $\int_a^b h f(h) dh$ when evaluated at s^* , it follows that the second term in braces in (A3') is positive at s^* . And since the first term in (A3') is clearly positive at s^* , (A3') must be positive.

We remark also that it is obvious that (10) will hold if q and y are sufficiently high, for as q approaches 1 and y approaches b , $x_\ell(b)$ approaches $x^*(b) > s^*$.

(v) If $s^{**} = s^*$, then no party chooses care exceeding s^* : Since (10) implies that $s^{**} < s^*$, we know that in this case $x_\ell(b) \leq s^*$. Our claim therefore follows.

(vi) If $x_\ell(b)$ is sufficiently low, then $s^{**} = s^*$: Assume the contrary. Then in particular it must be possible

that $s^{**} < s^*$ for an $x_\ell(b) \leq x^*(a)$. But by (ii) we know that if $s^{**} < s^*$, then $x_\ell(b) > s^{**}$. Hence, $x^*(a) \geq s^{**}$. This, however, contradicts (ii), so certainly for all $x_\ell(b)$ as low as $x^*(a)$, $s^{**} = s^*$.

We remark also that it is clear that as q decreases, so does $x_\ell(b)$, and it approaches 0 as q approaches 0; and similarly, as y decreases and approaches a . Hence, if q or y is sufficiently small, $s^{**} = s^*$. Q.E.D.

Footnotes

* Professor of Law and Economics, Harvard Law School.
I wish to thank L. Bebchuk and A.M. Polinsky for comments
and the National Science Foundation (grant no. SES-8014208)
for financial support.

An earlier version of present paper is contained in my
"Harm as a Prerequisite for Liability", Harvard University,
1979, and an elaborated, non-technical version of the present
paper is presented in Shavell (1984).

1. See the brief discussion in the concluding comments.

2. It will be clear that the analysis would not be
changed in an essential way were the possibility considered
that enforcement of the regulatory standard was not perfect
or was probabilistic -- and thus that some parties might be
able to engage in their activities without adhering to the
regulatory standard.

3. This comparison of liability and regulation may be
contrasted with the comparison made in the interesting and
well known paper of Martin Weitzman (1974) by considering
the problem of pollution damage. In respect to this problem,
Weitzman compared regulation of the amount of pollution to
use of a pollution tax, and he emphasized that there is no
natural advantage of either: not only might a social authority

err in deciding on a regulatory standard owing to lack of information about the costs and benefits of preventing pollution, it might err also in setting a pollution tax, for it would not generally be expected to know ahead of time the amount of damage that would be done by pollution.

Here, on the other hand, it is not use of a pollution tax that is compared to regulation, it is liability. The significance of this difference is that whereas the pollution tax is imposed before harm occurs and therefore naturally involves uncertainty as to its amount, liability is imposed by its nature only when harm has occurred, and therefore presumably involves little (and in the model, no) uncertainty over its amount. In other words, the problem here concerning liability as an alternative to regulation does not involve lack of information on the part of the social authority; instead, as explained, it has to do with dilution of incentives.

4. As will be seen, the assumption is really that not only does the regulator alone "observe" the values of the parameter, but also that the regulator does not communicate its value to the parties. The justification for the assumption is that sometimes information about risk cannot easily be conveyed to parties because of its complexity or its technical character.

5. We assume that the (2) holds for the first-best x for all h in $[a,b]$, that is, that $1 < -p'(0)a$.

6. Differentiate (2) with respect to h to get $0 = -p''(x)x'(h)h - p'(x)$, so that $x'(h) = -p'(x)/(p''(x)h) > 0$ (since $p'(x) < 0$ and $p''(x) > 0$).

7. In saying here that a party is liable for harm done, we are implicitly assuming that the form of liability is strict. Were we to consider instead the negligence rule -- under which a party would be liable only if his level of care was inadequate, that is, less than $x^*(h)$ -- the qualitative character of the results to be obtained would be essentially unaltered. For instance, referring to the next Proposition, the incentive to take care would still generally be diluted if $q < 1$ or $y < h$; it is only that $x_2(h)$ would be determined by a slightly different equation. (Actually, $x_2(h)$ would equal $x^*(\min\{h, y\})$ unless q and y were sufficiently high, in which case it would equal $x^*(h)$.)

8. Here $h(\)$ denotes the inverse of $x_2(\)$ where the latter is rising; thus $h(x_2(h)) = h$ for h in $[a, y]$ if $y < b$ and for all h in $[a, b]$ otherwise.

9. A solution to (8) and (8') exists since (8) is continuous in s and it will be shown in step (i) of the proof that s may be taken to lie in a closed and bounded interval.

10. It would be tedious to consider formally here (and later) the cases where $q < 1$ and $y \geq b$ or where $q = 1$ and $y < b$. But it will be obvious from the proof that in the former case, (a) and (b) will still hold. And a slight modification of the proof can be used to show that if $q = 1$ and y is sufficiently high, then $s^{**} = 0$ (the analogue to (a)), and that otherwise $s^{**} = s^*$.

11. Or, more precisely, s^{**} could equal any $s \leq x^*(a)$, for use of such an s would not affect any party's behavior.

12. Condition (12) is really the analog of (6), as is apparent when the latter is rewritten as $1 = -p'(s)E(h)$.

13. Note that although in the figure the graph of x^* lies below that of x_ℓ when $E(h_2) > h_2$, this need not be the case (e.g., if h_2 were less than but sufficiently close to $E(h_2)$, then the graph of x^* would be above that of x_ℓ).

14. In (19) the function $h_1(\)$ denotes the inverse of $x_\ell(\)$, and thus is analogous to the function $h(\)$ of the previous section (see note 8 above).

15. Equally, $s^{**}(h_2)$ could equal any $s \leq x_\ell(a)$, for use of such an s would not affect any party's behavior.

16. Since $p(x) = e^{-.1x}$, the first-order condition (2) is $1 + .1e^{-.1x}h$, so that $x^*(h) = 10\log(.1h)$; the graph is thus a plot of $10\log(.1h)$ for h in between \$10,000 and \$100,000.

17. Under liability, care taken is, by (4) and note 16, $x^*(q\min\{h,y\}) = 10\log(.1q\min\{h,y\})$; the lower graphs show the values of this function of h for three different sets of assumptions about q and y .

18. From (6) and note 16, we have $s^* = x^*(E(h)) = 10\log(.1(55,000)) = 86.12$.

19. From note 16 we have $x^*(h) = 10\log(.1(h_1 + h_2))$.

20. In this case we have from (15) and note 16 that $x_2(h_1) = 10\log((.1)(.9)E(\min\{h_1 + h_2, 80,000\}))$, where $E(\min\{h_1 + h_2, 80,000\}) = (1/3)\min\{h_1, 80,000\} + (1/3)\min\{h_1 + 10,000, 80,000\} + (1/3)\min\{h_1 + 20,000, 80,000\}$. (Consideration of only this one of the three cases will be enough to illustrate the main points of the extension of the model, as they concern parties' perception of risk.)

21. We use the word "estimate" here even though in fact the parties do not employ in their calculations a point estimate of h_2 but rather its probability distribution.

22. From (7) and note 16, we have $s^*(h_2) = x^*(E(h_1) + h_2) = 10\log(.1(45,000 + h_2))$.

23. This is the theme that is developed in Shavell (1983).

24. It should be noted in this regard that widespread ownership of liability insurance reduces the likelihood that a party would be unable to pay for harm done (and although it might also alter incentives to take care, this can be argued not to provide an independent reason for regulation).

25. In terms of the model, this paragraph may be summarized by saying that if we had to make a broad generalization concerning the area of the typical tort, it would be that parties possess superior information, so that the basic version of the model applies, and also that y and q are both fairly high.

26. This paragraph may be summarized by saying that in the areas where regulation is significant and the basic model applies, the informational advantage of the parties might not be great; that where the extension of the model applies, the regulator may have an informational advantage about an important aspect of risk; and that y or q may be low.

27. See Prosser (1971), pp. 203-204; thus use of regulation is indeed accompanied by use of liability.

28. See Proposition 4.

29. See George Stigler (1971), Richard Posner (1974), and Sam Peltzman (1976).

30. In the case of the numerical example of the basic model (Section III), the likelihood of an accident under liability is on average about .0003. Thus, expected administrative costs under liability will be lower than those under regulation unless the cost of determining adherence to regulation is much smaller than the administrative cost of the liability system (which should not be equated with the costs of going to court, for the parties to disputes settle far more often than not). For instance, suppose that the administrative costs of the liability system is \$200 per accident. (If, say, 90% of disputes are settled at an administrative cost of about \$100 and 10% go to trial at an administrative cost of \$1000, the expected administrative cost given that an accident has occurred would be approximately \$200). Then the expected administrative costs associated with use of liability would be only $.0003 \times \$200 = \0.18 . Hence the cost of determining adherence to regulation would have to be less than this amount for regulation to involve lower administrative costs. While this might seem

unlikely, it must be admitted that significant savings in administrative costs might be achievable through use of probabilistic means of enforcement of regulation.

31. See Wittman (1977), which discusses closely related issues.

32. For any h such that $x^*(h)$ lies in the domain of the function $h(\cdot)$, we have $h(x^*(h)) > h$; this is obvious from the graph in Figure 4.

References

Sam Peltzman, "Toward a More General Theory of Regulation," Journal of Law and Economics, 19 1976, 211-240.

Richard Posner, "Theories of Economic Regulation," Bell Journal of Economics, 5 Autumn 1974, 335-358.

William Prosser, Handbook of the Law of Torts, 4th ed., 1971.

Steven Shavell, "Liability for Harm versus Regulation of Safety," forthcoming in Journal of Legal Studies, 1984.

George Stigler, "The Economic Theory of Regulation," Bell Journal of Economics, 2 Spring 1971, 3-21.

Martin Weitzman, "Prices vs. Quantities," Review of Economic Studies, 41 1974, 447-491.

Donald Wittman, "Prior Regulation versus Post Liability: The Choice Between Input and Output Monitoring," Journal of Legal Studies, 6 1977, 193-212.

</ref_section>